



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

Grice for graphics: Pragmatic implicature in network diagrams

Citation for published version:

Oberlander, J 1995, 'Grice for graphics: Pragmatic implicature in network diagrams', *Information Design Journal*, vol. 8, no. 2, pp. 163-179. <https://doi.org/10.1075/idj.8.2.05obe>

Digital Object Identifier (DOI):

[10.1075/idj.8.2.05obe](https://doi.org/10.1075/idj.8.2.05obe)

Link:

[Link to publication record in Edinburgh Research Explorer](#)

Document Version:

Early version, also known as pre-print

Published In:

Information Design Journal

General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.



Grice for graphics: pragmatic implicature in network diagrams*

Jon Oberlander[†]

Human Communication Research Centre
University of Edinburgh
2 Buccleuch Place, Edinburgh, EH8 9LW

1 Introduction

Petre and Green [1992] observe intriguing phenomena which occur when real designers use complex graphical representations. They also offer a promising account for these phenomena. This paper explores an alternative account, which relies on the notion of *graphical implicature*. Whether or not the current account is useful has broader significance, because if it is useful, it helps supply the first premise for the following argument: (i) there are certain parallels between pragmatic phenomena in natural language (NL) and graphical representation; (ii) formal techniques have been developed for modelling some of the NL phenomena; and thus (iii) once the graphical data are better-understood, we may be able to treat them with formal techniques from NL pragmatics.

The paper has the following structure. First, we introduce the phenomena discussed by Petre and Green, and summarise their account. Then we introduce some key concepts from linguistic pragmatics, focussing on Grice's [1975] theory of implicature. We then indicate how Grice's ideas have been applied in the graphical domain, by Marks and Reiter [1990]. Finally, we show how the ideas apply to Petre and Green's observations, and relate the two accounts.

2 Network diagrams

Petre and Green [1992] report on expert users' use of graphical notations for computer-assisted design in electronics (CAD-E). They interviewed five engineers, all with over ten years' ex-

*Based on a talk presented at the Summer School organised by the Joint Councils' Initiative in Cognitive Science and Human-Computer Interaction, held at the University of Edinburgh in September 1993. Thanks to Corin Gurr, John Lee and Keith Stenning for helpful comments; I am particularly grateful to Thomas Green for his encouragement, without which this paper would never have been started, let alone finished.

[†]EPSRC Advanced Fellow at the Human Communication Research Centre. The support of the Economic and Social Research Council for HCRC is gratefully acknowledged.

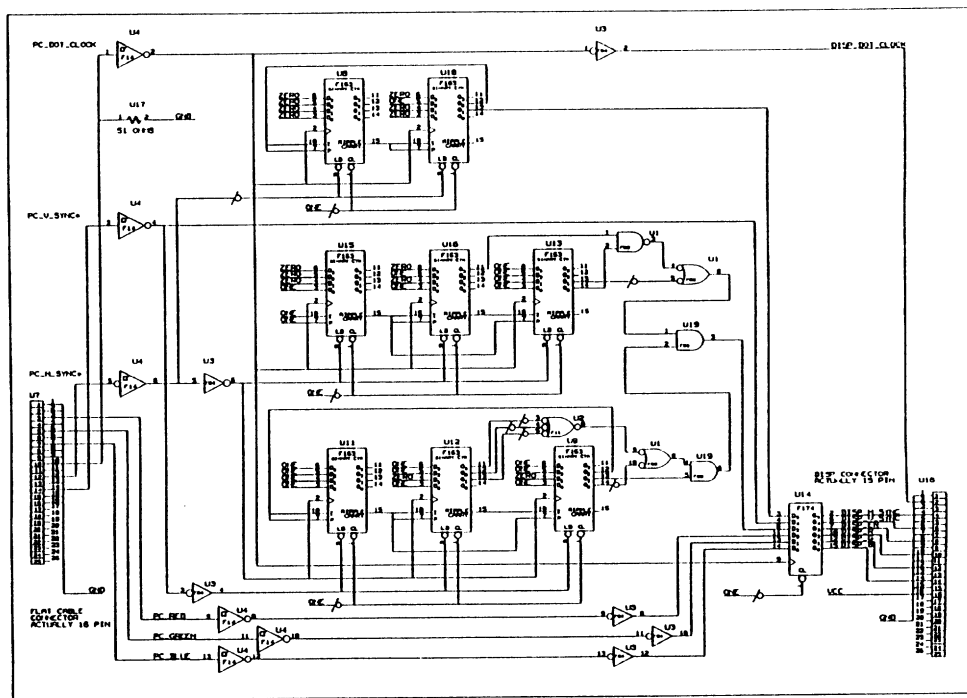


Figure 1: A schematic drawing. Reproduced from Petre and Green's Figure 1 [p52].

perience in digital electronic design, and regular users of CAD tools. The engineers demonstrated their tools and activities, and discussed their practices. The key representation for such users is the *schematic drawing*, an example of which appears in Figure 1. Such drawings are largely graphical, and the CAD tools permit multiple layering, allowing zooming in and out. The drawings are also supported by textual annotations and appendices.

2.1 Petre and Green's observations

Petre and Green [1992:51–65] report seven themes that emerged in their interviews. Their main points are reproduced verbatim here:

Overviews The designers found graphics better for overviews [p51].

Zooming Designers emphasize the need for improved detail on the overall-gestalt-like views obtained by changing scale, i.e., zooming [p54].

Neighbourhood: adjacency and locality The nature of graphics allows *connectedness* to be represented by a line, and thus kept separate from *relatedness* (represented by adjacency). Hence, adjacency or position is always, even locally, a secondary cue, a manipulable variable, available for reinforcing associations, for suggesting structure, or for giving hints to the reader. Text, however, must either use adjacency to indicate *both* relationships, or introduce symbolic links (e.g., cross-references or variables) [p55].

Shifts from graphics to text Text has its uses; even after reasonable evolution, electronics schematics remain a multi-level system with alternative representations in text [p57].

Viscosity Massive changes to graphic representations are hard work: if a group of components is moved, connections have to be re-established. Similarly, the cases where moving a piece of text is difficult are where connectivity has been established explicitly; in other cases, blocks of text may be moved freely without difficulty [p61].

Search trails Searching is well supported in CAD-E in the domain of connectivity, but not in the domain of functionality [p63].

Vocabulary and space consumption Although it is commonly believed that graphics is more compact than text, the comparison is not straightforward, particularly because most graphical notations rely to some extent on text to keep their vocabularies manageable. *Complete* system descriptions tend to be massive whether in graphics or text [p64].

2.2 An explanation in terms of cognitive dimensions

Their primary claim is that these observations are to be expected on the basis of Fitter and Green's [1979] criteria for information accessibility; or from Green's [1989] cognitive dimensions, some of which refine the Fitter and Green criteria.

Fitter and Green state five criteria, whose satisfaction contributes to information accessibility in a given diagrammatic notation: *relevant* information should be presented in a perceptual form; *redundant recoding* should re-present important symbolic information in diagrammatic form; *restrictions* should be placed upon users, so that only easily understood objects are presented; underlying mechanisms should be directly *revealed*, and *respond* to manipulation; *revision* of the diagram should be easy.

Green uses the notion of cognitive dimensions to characterise notations in the context of their support environment, and their users' tasks. *Viscosity* is a measure of resistance to local change; *hidden dependencies* are frequent when structural relationships are not made explicit in the notation; *role-expressiveness* measures the ease of seeing the function of substructures in the notation; *premature commitment* occurs when users are forced into making decisions too early in a process.

Petre and Green [1992:50] observe that two other predictions flow from these models. First, that users need to *escape from formalism*, so that if a task requires information to be stored, and that information cannot be captured in the chosen notation, then users will switch to another notation to achieve their goal. Secondly, users will adopt *secondary perceptual notations* to capture or emphasise certain information. That is, they will go beyond the "official" interpretation of the diagrammatic symbols, and—for instance—use spatial arrangement to encode useful information. In fact, they later concede [p66] that the Fitter and Green model does not really predict the use of secondary notation, because it focusses exclusively on the intended interpretation of the notation; and nor does the model deal with the escape from formalism, because it focusses on the properties of single notations, rather than seeing them as forming elements in a set of tools. In addition, constraints on vocabulary size are not really accounted for by the model, and it does not address differences among individuals' skill with the notations.

They argue that cognitive dimensions provide a more useful framework, since a process model forces us to provide a mechanism for the use of notation, in a rich context, embracing both tasks and tools.

2.3 An explanation in terms of Gricean pragmatics

We agree that cognitive dimensions offer a useful starting point for explanations, but present here another approach, which we believe to be complementary. This approach, exploiting the notion of *graphical implicature*, makes for stronger explanations in some areas than cognitive dimensions, but is correspondingly weaker elsewhere.

The next sections provide some background on linguistic pragmatics, and a description of previous work on graphical implicature, before we return to discuss Petre and Green's data.

3 What is pragmatics?

According to Morris [1938], the study of systems of communication can be divided into three parts. Syntax is devoted to 'the formal relation of signs to one another'; semantics to 'the relations of signs to the objects to which the signs are applicable'; and pragmatics studies 'the relations of signs to interpreters'. Although the notion of 'signs' can be read very widely, most work in pragmatics has followed in the footsteps of syntax and semantics, and focussed on language use. Over time, differing pragmatic traditions have emerged. Followers of Austin [1962] emphasise that language use is a form of action, and investigate the conceptual relations among *speech acts*, their preconditions and effects (cf. for example Searle [1969]). By contrast, work in the Gricean tradition emphasises the view that communication takes place by virtue of a set of defeasible assumptions about rational co-operation (Grice [1975, 1978]). A key Gricean idea is that a participant in a discourse will understand utterances directed at her in part because she recognises that the utterances were constructed with the intention that she so recognise them. Griceans have thus extensively investigated the linguistic role of defeasible inference, or *implicature*, particularly where this concerns mental states.

The divergence of these schools of thought has contributed to uncertainty over the scope of pragmatics, and the 1970s saw various attempts establish its boundaries. Stalnaker [1972] attempted an extensional definition, and held that pragmatics included the study of deixis (at least in part), implicature, presupposition, speech acts, and aspects of discourse structure. Thomason [1977] maintained that a proper account of inference is crucial to a pragmatic theory, and that working out how language and context interact is the major problem for an account of language understanding. Gazdar [1979] proposed the influential equation 'Pragmatics = Meaning – Truth Conditions' ($P=M-TC$). On this basis, pragmatics is devoted to the study of all the aspects of meaning which lie outside the mere conditions of truth for sentences. For Gazdar, one of the most important aspects of an utterance's meaning is the way it *changes the context* in which it is uttered. Levinson [1983:14] concurs with Gazdar, and amplifies the meaning of $P=M-TC$, by providing an inventory of an utterance's communicative content. The list includes: truth conditions or entailments; conventional implicatures; presuppositions; felicity conditions; conversational implicatures, both generalised and particu-

larized; and inferences based on discourse structure.

Of these, implicatures are perhaps the most important; we now turn to Grice's own theory of implicature.

4 The Gricean Theory of Implicature

First, let us give two concrete examples of what Grice takes to be implicature. Consider the dialogue in (1), and the alternative texts in (2):¹

- (1) a. A: Can you tell me the time?
- b. B: Well, the milkman has come.
- (2) a. The lone ranger jumped on his horse and rode into the sunset.
- b. The lone ranger rode into the sunset and jumped on his horse.

(1a) has the implicature that *A* would like *B* to tell her the current time. (1b) carries the implicature that *B* does not know the precise time, but has some information which may permit *A* to compute it. (2a)'s implicature is that the jump happened first, and was followed by riding. By contrast, (2b)'s implicature is that the riding preceded the jumping. In both (1) and (2), implicatures go *beyond* the literal truth conditional meaning. For instance, all that matters for the truth of a complex sentence of the form *A and B* is that both *A* and *B* be true: the order of mention of the components is irrelevant. As $P=M-TC$ suggests, implicatures help to bridge the gap between truth conditions and "real" meaning.

But how do implicatures actually function in communication? Grice's account is deceptively simple. His claim is that conversational participants adhere to a *Co-operative Principle* (CP). This can be phrased as an injunction: make your contribution such as is required, at the stage at which it occurs, by the accepted purpose or direction of the talk exchange in which you are engaged (Grice [1975:26]).² In itself, this is very general, and Grice suggests four maxims, which can be thought of as guiding usual communication: Quality, Quantity, Relation and Manner. These are spelt out in Figure 2.

How does an individual *H* in a conversation actually retrieve the implicatures *q* carried by an utterance by a speaker *S*? Grice suggests that the following pattern of inference is typical. *H* starts from the premise that *S* has said that *p*, and notes that there's no reason to think *S* is not observing the maxims, or at least the CP. *H* then reasons that in order for *S* to say that *p* and indeed be observing the maxims (or the CP), *S* must think that *q*. *S* must also know that it is mutual knowledge that *q* must be supposed if *S* is to be taken to be co-operating. Now, *S* has done nothing to *H*, the addressee, thinking that *q*. So *H* concludes that *S* intends *H* to think that *q*, and in saying that *p* has implicated *q*. There is, in principle, no reason to think that individuals actually pass through this cycle of inference every time they work out what a speaker means. But, for at least one class of implicatures, the pattern can be reconstructed if necessary.

¹The linguistic examples, (1)–(7) are all taken directly from Levinson [1983].

²Page references to Grice are to versions in Grice [1989].

The maxim of Quality Try to make your contribution one that is true, specifically:

1. do not say what you believe to be false
2. do not say that for which you lack adequate evidence

The maxim of Quantity

1. make your contribution as informative as is required
2. do not make your contribution more informative than is required

The maxim of Relation Make your contributions relevant

The maxim of Manner Be perspicuous, and specifically:

1. avoid obscurity
2. avoid ambiguity
3. be brief
4. be orderly

Figure 2: Grice's conversational maxims. The four maxims can be thought of as guiding behaviour so as to comply with the Co-operative Principle. The subcomponents of a maxim—such as the four parts of the maxim of manner—can be termed *submaxims*. From Grice [1975:26–7].

4.1 The family of implicatures

Grice carefully categorises implicatures into various types. First, he contrasts *conventional* with *non-conventional*. Conventional implicatures are inferences attached to individual lexical items, which are not licensed directly by conversational maxims. Connectives such as *but* and *therefore* are of this type, and the pattern of inference just outlined does not hold for them. Consider (a):

- a.
- b. He is an Englishman; he is, therefore, brave.
- c. *His Englishness implies his braveness*

According to Grice [1975:25], an utterer of (aa) has committed themselves to the truth of both the antecedent and consequent. However, although they have certainly implied that braveness flows from Englishness, as in (ab), they have not “strictly” said it, so it can only be an implicature—(ab) could be false while (aa) is true. By contrast, the pattern of reasoning governed by CP does work for non-conventional implicatures. Among these, there are two classes: *non-conversational* and *conversational*. Of the latter, there are two sub-classes: *generalised* and *particularised*.

- (3)
 - a. I walked into a house.
 - b. *The house \neq my house.*
- (4)
 - a. A: What on earth has happened to the roast beef?
 - b. B: The dog is looking happy.
 - c. *Perhaps the dog has eaten the beef.*

(3) represents a generalised implicature; the inference from (3a) to (3b) goes through, with no special context required. By contrast, (4) represents a particularised implicature, where the inference goes through only in a particular context; here (4b) only implicates (4c) in the context of (4a).

4.2 Observing, violating and flouting

Conversation proceeds smoothly so long as people *observe* the maxims. Things go wrong when people *violate* them. However, Grice is well aware that people don't simply follow the conversational maxims in a flat-footed way. We are perfectly capable of exploiting them for effect: we can disobey, or *flout* a maxim or set of maxims. Take the maxim of quantity, and consider (5) and (6).

- (5)
 - a. Nigel has fourteen children.
 - b. *He has exactly fourteen children.*
- (6)
 - a. Either John will come, or he won't.
 - b. *There's no point in worrying about it.*

A straightforward utterance of (5a) implicates (5b): saying that someone has fourteen children when they actually have fifteen would be true, but would violate the injunction to be as informative as is required. By contrast, an utterance (6a) appears to flout the maxim, since it is completely uninformative, and merely states a tautology. Someone who hears (6a), and believes that its utterer is still observing the Co-operative Principle, will have to work out how voicing this tautology could add to the conversation. One solution is to infer that the speaker means that there is a fact of the matter, but that this is inaccessible to the conversational participants; hence (6b) is a natural implicature.

The distinction between observing and flouting can also be illustrated by considering the maxim of manner. Take (2) and (7):

- (2)
 - a. The lone ranger jumped on his horse and rode into the sunset.
 - b. *First he jumped, and then he rode.*
- (7)
 - a. Miss S produced a series of sounds corresponding closely to the score of an aria from *Rigoletto*.
 - b. *She wasn't singing in the conventional sense.*

An utterance of (2a) usually implicates (2b), because speakers observe the injunction to be orderly, and describe events in the same order in which they happened. By contrast, the utterance (7a) flouts the injunction to be brief. In interpreting the utterance, a hearer would normally conclude that the speaker is describing Miss S as singing. Since this much shorter description was not used, the hearer who believes that the Co-operative Principle is still operating will have to find a reason why not. One possibility is that (7b) holds, so that (by the maxim of quality), the speaker could not have uttered a briefer description.

This brief summary of the Gricean theory of implicature should be sufficient for current purposes. One issue which we have not yet touched on is how to tell when a proposition is (say) a conversational, rather than conventional, implicature. This is an issue to which we return in section 5.4.

5 Implicatures in automatic network diagram design

Marks and Reiter [1990] explored the possibility of applying the Gricean theory of implicature to graphical representations. In particular, they addressed the automatic design of network diagrams, and developed their system, ANDD to avoid unwanted conversational implicatures. Here, we summarise ANDD's task and structure, and then note the types of graphical implicatures which Marks and Reiter maintain should be avoided. We conclude this section by discussing their arguments for three additional conversational maxims.

5.1 ANDD

ANDD's task is to design network diagrams which accurately portray input network models. The domain of the models could be, for instance, configurations of clients and servers on a computer network; or they could be components in an electronic circuit. ANDD's network diagrams have both syntax and semantics. The *morphological elements* are node symbols, link symbols, text labels, and diacritical symbols. Each class of symbols has a set of *graphical properties*: for instance, node symbols have shape, size, pen colour, and fill colour. The syntax also specifies a set of *perceptual organisation relations*, comprising: sequential layout (top-bottom; left-right), proximity grouping, alignment, symmetry, similarity, and ordering. Semantically, an *expressive mapping* relates the network-model to the syntax of the network-diagram: for instance, vertices and edges are related to node and link symbols in Figure 3. ANDD generates a diagram by using three rule-based systems to create an expressive mapping, choose suitable values for graphical properties, and choose suitable locations for the symbols. The last stage is computationally hard, so ANDD uses syntactic constraints together with existing local layout to heuristically generate a full layout.

In fact, Marks and Reiter's claim is that many possible layouts are pragmatically inappropriate, and that if the system can detect unwanted graphical implicatures, avoiding them will help further constrain the space of full layouts.

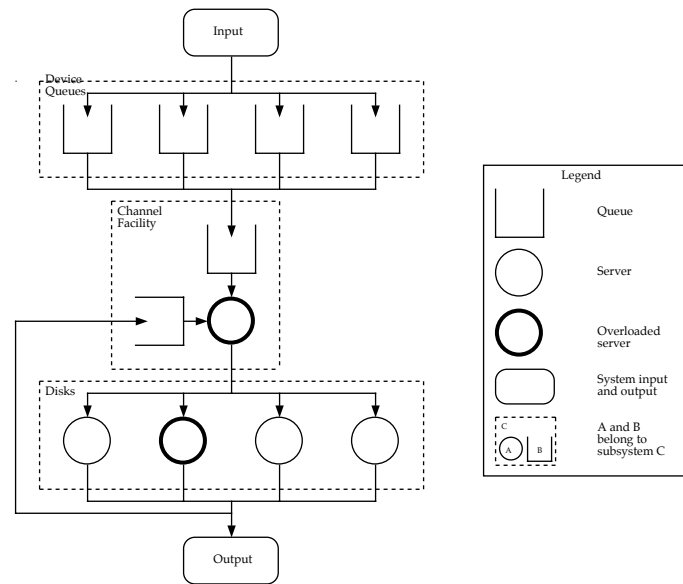


Figure 3: A network diagram of a computer's disk subsystem. This diagram has no unwanted implicatures. Based on Marks and Reiter's Figure 1 [p456].

5.2 Avoiding unwanted implicatures

Assuming a particular network model, we can stipulate that the network diagram in Figure 3 conveys precisely the intended information. Given this, Marks and Reiter maintain that the network diagrams in Figures 4 and 5 carry unwanted conversational implicatures.

They suggest that Figure 4 is misleading for at least four reasons. First, the pen-width used for the channel-facility queue differs from pen-width for all other queues. The implicature is that this queue must be different in some respect, since otherwise the same pen-width would have been used. Secondly, disk symbols are perceptually grouped into two gestalts. The implicature is that there is some basis in the domain for this grouping, since otherwise elements would have been placed regularly. Thirdly, device-queue symbols are laid out irregularly. The implicature is that one of them is uniquely different, since otherwise elements would have been placed regularly. Finally, a different font is used for the channel-facility's text label. The implicature is that it must have different sub-system status, since otherwise the same font would have been used.

Note that the diagram in Figure 4 uses the same expressive mapping as that in Figure 3, and that all the features of the assumed network model are therefore faithfully portrayed in both diagrams. The difficulty with Figure 4's diagram is not that it is *incorrect*, but that various aspects of it—not specified by the expressive mapping—are misleading. All four of Marks and Reiter's problems arise because there is a general expectation of *regularity*, and any departures from this are assumed to be meaningful: the diagram contains irregularities where the domain has none. In one case, involving the grouping of disk symbols, an irregularity gives rise to a further regularity, so the diagram also contains a regularity where the domain has none. Figure 5 provides a further example of an "over-regular" diagram. Its node symbols are ordered by size, and the implicature is that there must be a similar ordering relation

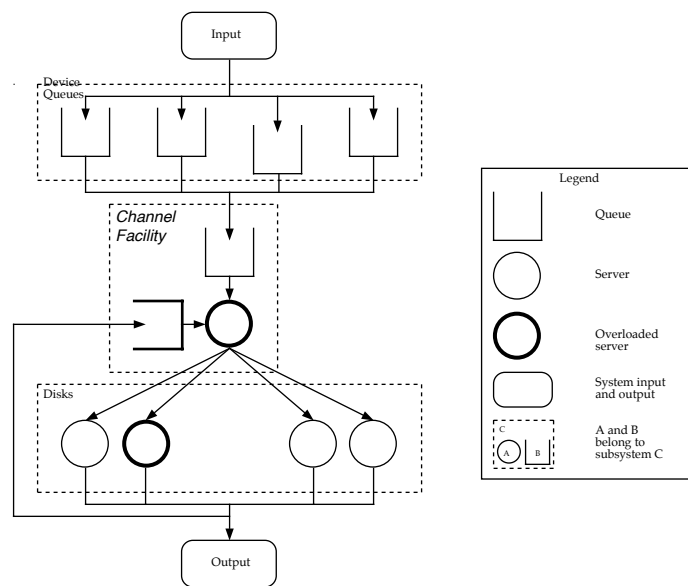


Figure 4: A variant on Figure 3, possessing unwanted implicatures. Based on Marks and Reiter's Figure 2 [p456].

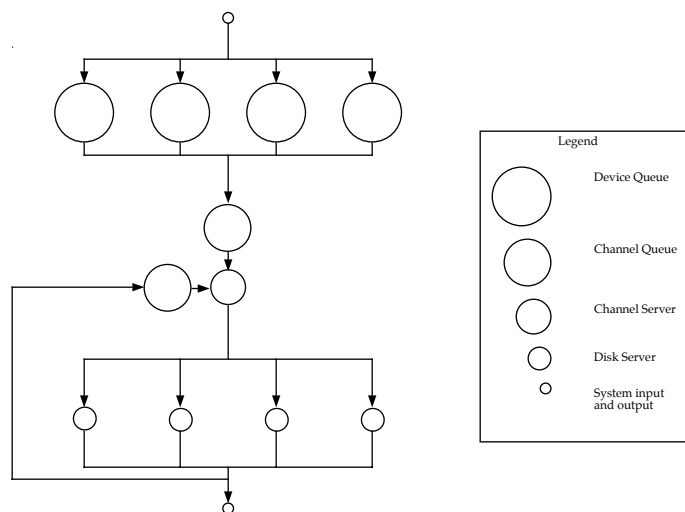


Figure 5: Another variant on Figure 3, possessing different unwanted implicatures. Based on Marks and Reiter's Figure 3 [p456].

among the vertices in the network model. Assuming that we are still dealing with the same network model, this implicature is again misleading.

So, it is claimed that Figures 4 and 5 are defective for pragmatic, rather than semantic reasons. That is: they are correct with respect to the expressive mapping, but give rise to spurious implicatures. But if a network diagram's conversational implicatures are the culprit, it must be possible to sketch a Gricean account which cites the conversational maxims which are being violated.

Marks and Reiter attempt to do just this. For instance, they suggest that unwanted implicatures arise if two graphical properties (such as shape and colour) are used to communicate one network attribute, since the independence of the properties will spuriously suggest that other shape/colour combinations might exist. Such redundancy runs counter to the maxim of quantity, since in some sense the diagram's contribution is more informative than is required. However, they also note that redundancy—or more generally, extra information—is actually often required. For example, Figure 3 contains background context to allow the role of the overloaded servers to be understood; it also uses perceptual grouping as well as diacritical enclosure to convey sub-system affiliation.

Returning to the faults in Figure 4, they observe that the use of spurious graphical-property values may do their damage by violating the maxim of relation. What is wrong with some of the departures from regularity, such as the distinct pen-width for the channel-facility queue, is that irrelevant properties are introduced. As Marks and Reiter observe, such problems may equally well be attributable to violations of the maxim of manner. Certainly, other departures from regularity are more easily explained in this way: for example, the irregular layout of the device-queue symbols violates the injunction to be orderly.

5.3 Elaborating the maxims

Marks and Reiter propose that the submaxims of manner must be augmented with additional concepts to deal with the graphical domain. In particular, a diagram generator must strive for appropriate perceptual organisation, and observe human perceptual limits. Spurious perceptual organisations that are orthogonal to the diagram's intended information should be avoided. Graphical-property values should be limited in number, so that they are easily distinguished, and perceptually dissimilar.

These are useful notions, although it is possible to argue that they overlap to some extent with Grice's original submaxims of manner. Spurious perceptual organisations take two forms: diagrams with too little regularity with respect to the domain; and diagrams with too much regularity. Grice's injunction to be orderly is violated by the first form. The additional concept that is required is essentially a dual—don't be *too* orderly. We could therefore break the orderliness submaxim in two, in the manner of the maxims of quality and quantity, as in Figure 6. Perceptual limits are perhaps more controversial. Nonetheless, it might be argued that the injunction to observe perceptual limits is, in fact, another way of expressing the injunction to avoid ambiguity. If I find it difficult to distinguish two graphical-property values, a given value may be regarded as ambiguous, denoting one or other network attribute.

Marks and Reiter also follow Hirschberg [1991] in suggesting that the failure to use basic-level

Submaxim of Order

1. make your contribution as orderly as is required
2. do not make your contribution more orderly than is required

Figure 6: A revised version of order, a submaxim of the maxim of manner.

predicates (after Rosch [1978]) can be a further source of unwanted implicatures. In natural language, a predicate such as *dog* is basic-level, while *animal* is superordinate, and *poodle* is subordinate. Departures from basic-level are obviously permitted, but always carry additional implicatures. Marks and Reiter suggest that the preference for basic-level expressions carries over to diagrams. For example, it is assumed that the basic-level, preferred status of symbols in a computer display is static and non-blinking. So if a diagram designer chooses blinking symbols for the whole display, the user will interpret the choice of non-basic level attributes as meaningful. Again, this is a useful suggestion, although Fodor [1983:96] argues that the use of basic-level expressions is in fact already dictated by Grice's maxim of quantity and submaxim of brevity. Basic-level predicates provide a crucial layer in language, where brevity and informativeness are most effectively combined. So unprincipled choice of non-basic symbols violates either the injunction to be brief, or the injunction to be informative.

5.4 Tests for implicatures

Marks and Reiter's discussion provides useful examples of graphical implicatures. But is there a general way of telling when part of a diagram's meaning has been generated by an implicature? And is there a way of telling which kind of implicature—conventional or conversational—is responsible? Within the linguistic domain, Grice proposed a number of tests for conversational implicature. The precise set has been the subject of argument, but Levinson's [1983:119–20] six properties are a good summary; the first four are from Grice, the fifth from Sadock [1978], and the last is Levinson's own:

Cancellability (or defeasibility) If extra information is added to the context, an implicature that normally holds can be over-ridden.

Non-detachability The precise surface form of the expression is irrelevant, so the implicature will go through if another expression with the same meaning is substituted for it.

Calculability The implicature can, in principle, be computed from first principles, via the Co-operative Principle and background knowledge.

Non-conventionality Implicatures are not part of the conventional meaning of expressions.

Reinforceability We can make the implicature explicit without causing anomalous redundancy.

Universality If the implicature is generated via an assumption of co-operation, all languages should share it.

None of these is entirely uncontroversial, but they can all be defended against the most obvious criticisms. For our current purposes, we should note three points. First, both defeasibility and non-conventionality help to distinguish implicatures from truth-conditional meaning. Secondly, there is a trade-off between calculability and conventionality: something which is calculable will usually be non-conventional; something which is conventional (such as the contrastive sense to *but*) will usually not be calculable. Finally, the principle of non-detachability does not govern implicatures generated by the maxim of manner, since this is clearly concerned with the surface form of conversational contributions.

We can now assess whether the graphical implicatures discussed by Marks and Reiter are genuinely conversational in nature.

They are certainly defeasible in general: textual annotation can generally over-ride any graphical implicature. For instance, the implicatures associated with font choice can be defeated by adding text to the diagram's legend, such as *Choice of fonts is not meaningful*. In network diagrams, relative size of symbols does not carry implicatures concerning the size of domain objects, but in other domains, a frequent textual annotation is *Not to scale*, which cancels the relevant implicatures. Graphical implicatures can also be over-ridden directly by graphical means. In network diagrams, the submaxim of order leads to related components being grouped together, and unrelated components being separated. However, sometimes related components must be separated for incidental reasons. In these cases, the explicit lines connecting them defeat implicatures concerning unrelatedness. The defeasibility of graphical implicatures helps to show that they are conversational, and not conventional, since the latter cannot be over-ridden.

However, like conventional implicatures in language, many graphical implicatures are not non-detachable; this is, however, because they are traceable to the maxim of manner and its submaxims. Since these concern the surface form of representations, variations in the form can cause implicatures to vanish. The calculability of the implicatures has already been demonstrated, since we rehearsed the reasoning behind the implicatures relating to regularity and over-regularity. In theory, then, the non-conventionality of these implicatures is assured. However, it could be argued that some other graphical implicatures are actually conventional (and not calculable). For example, some complex graphical entities form patterns which are highly recognisable; the symbols associated with the bi-stable flip-flop discussed in section 6.1 are of this kind. Practised users of a notation will interpret such symbols "directly", without having to build up their interpretation from the meanings of their parts. These complex graphical entities thus resemble idiomatic phrases in natural language. However, although their meaning *can* be read off directly, it need not be. Thus, their interpretations, and associated implicatures, are still calculable.

The re-inforceability of graphical implicatures seems relatively uncontroversial; indeed, some of Marks and Reiter's problems with redundancy may be traced to reinforcement. When the diagram in Figure 3 uses perceptual grouping as well as diacritical enclosure to convey subsystem affiliation, the former works via implicature, but is reinforced via the semantically-specified enclosure symbolism. Universality may seem more difficult: it might be argued that the extraordinary range of different graphical notations means that there are few "universal" implicatures, carried by a symbol (or ordering of symbols) wherever they appear. But this is not necessary. What matters is that if a configuration has an implicature in one graphical

notation, the implicature should be retained by the configuration's translation in any other notation of equivalent expressiveness. The basic point is just that if the implicature arises from principles of co-operation, it should be reproducible in any equivalent notation used co-operatively. The universality property thus helps distinguish conventional implicatures from conversational. The real problem with using universality as a test will reside in the difficulty of establishing which notational systems are genuinely equivalent.

6 Implicatures in CAD-E

With these concepts in hand, we can re-examine Petre and Green's data on computer-assisted design in electronics. This section discusses a simple example in terms of graphical implicatures, and then assesses the extent to which an implicature account can explain the themes which arose from Petre and Green's interviews. Finally, we compare the implicature account with the one proposed by Petre and Green.

6.1 Design errors and unwanted implicatures

Looking at the schematic drawings, we can contrast those aspects of the network diagram which are fixed by the expressive mapping from those which are not. Only the former are part of what Petre and Green term 'the official notation' [p67]. We therefore have here a clear distinction between the semantics and pragmatics of diagrams—between the literal and the "full" meaning of the graphics. Let us compare the diagram in Figure 1 (repeated here as Figure 7) with that in Figure 8. According to Petre and Green, the two machines depicted are equivalent, but the one in Figure 8 fails to use what they term "manipulable cues". The diagram suffers from 'neglect of adjacency cues: things are grouped visually which are not related logically' [p57]; components are dispersed which ought to be related (for instance, U8 and U10); the signal flow is badly confused, via convoluted wiring; and two gates have been rendered to look like a bi-stable flip-flop, which they label 'a striking mis-use of a strong perceptual cue.' [p61].

In our current terms, the diagrams denote the same network model, and use the same expressive mapping. Where they differ is in their pragmatic status. The connections between the components are basically the same in the two diagrams, but the use of space is very different: the connections are specified in the network model, but the actual layout is a pragmatic matter. The diagram in Figure 7 observes the Gricean maxims, whereas that in Figure 8 violates them.

The specific faults in Figure 8 can be traced to the violation of particular conversational maxims. Since they are largely layout problems, they are best thought of as violations of the maxim of manner. When things are grouped visually which are functionally or logically distinct, the diagram violates the second part of our new submaxim of order (cf. Figure 6). It does so because the diagram immediately juxtaposes various symbols, and thereby generates the unwanted implicature that the symbols' denotations are in some way related. When related components are dispersed, the diagram violates the first part of the submaxim of order. By failing to juxtapose symbols which should be related, the diagram fails to be as orderly as

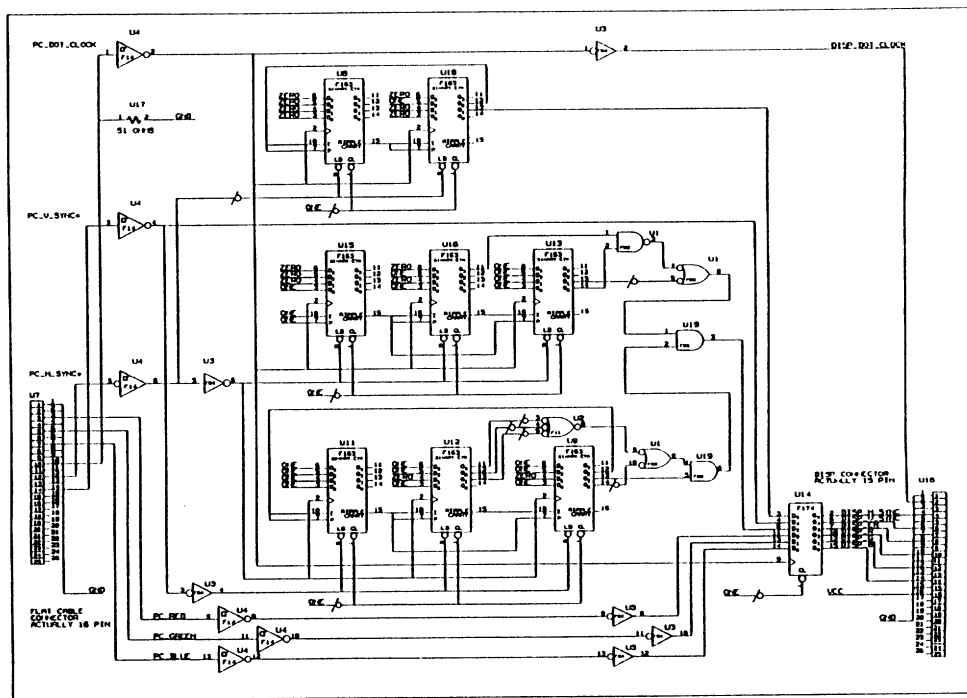


Figure 7: A competent electronic schematic drawing. Reproduced from Petre and Green's Figure 1 [p52].

it should. Convoluted wiring may be thought of as a violation of the submaxim of brevity, since connections are made longer and more complex than is justifiable.³ The fake bi-stable flip-flop can be thought of as violating the submaxim of ambiguity, because the gates' rendering leads naturally to an additional, unintended interpretation. Since this interpretation is also incorrect, the diagram may even appear to violate the maxim of quality—although it is strictly correct. The bi-stable flip-flop is particularly interesting, because it also raises the issue of diagram vocabulary and discriminability. We return to this issue towards the end of the next subsection.

6.2 Petre and Green's observations

So, the implicature-based account can help characterise specific design errors. But can it say anything about the emergence of the seven themes which characterise Petre and Green's interviews with CAD-E users? Let us consider each in turn.

Overviews 'The designers found graphics better for overviews.' A circuit graphic collects together many components in a single, synoptic representation. By contrast, to convey the same information through text requires the assembly of multiple lists and representations. The unary representation therefore juxtaposes many more components, and

³ Alternately, it may be maintained that unnecessary loops and angles are *irrelevant* contributions, and hence violate the maxim of relation.

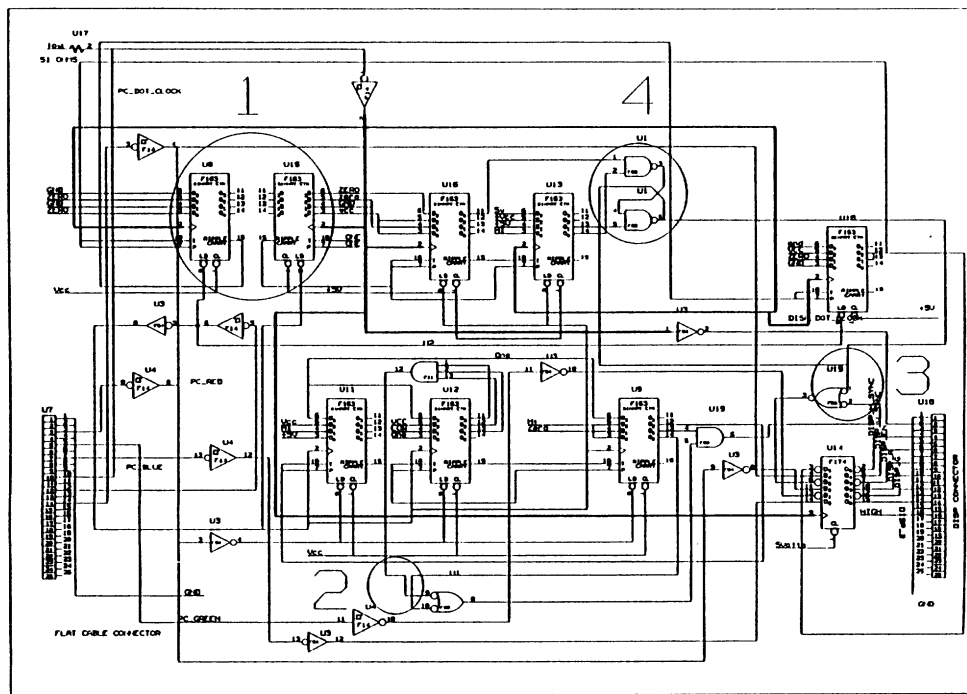


Figure 8: An incompetent electronic schematic drawing, representing the same circuit as in Figure 7. The circles enclose significant mis-cues: (1) symmetry misleadingly relates unrelated components; (2) extra bends in a wire have been left by past edits—there is also convoluted wiring at the left of the diagram; (3) Input 1 to the gate has redundant entry routes, and the angle of entry is odd; and (4) the gates misleadingly resemble a bi-stable flip-flop. Reproduced from Petre and Green’s Figure 4 [p56].

hence the scope of graphical implicatures is greater than the scope for textual implicatures in the alternative representations. It follows that the schematic drawing is a richer medium for pragmatically encoding information, such as functional relatedness. This may well be one reason why graphics often seem to say more than texts, even where these are known to be equivalent at the *semantic* level.

Zooming ‘Designers emphasize the need for improved detail . . . by changing scale.’ A synoptic representation restricts the granularity of representation. Zooming in on details via in-place expansion is particularly good because it retains graphical context, and therefore preserves the background’s implicatures, which specify implicit relationships.

Neighbourhood: adjacency and locality In schematic drawings, ‘adjacency or position is always, even locally, a secondary cue.’ The drawings separate out connectivity from relatedness, whereas text conflates them, or has to introduce symbolic linking. In fact, the main difference resides in the dimensionality of the modality: diagrams use two dimensions, and unformatted text uses only one. But both diagrams and texts carry implicatures relating to order. The “perceptual grouping” conventions governing the layout of diagram elements in Figure 7 play a similar role to the descriptive conventions governing the ordering of clauses in example (2). Textual ordering generalises to dia-

grammatical juxtaposition; but in both cases, extra information—beyond the semantic—is carried by pragmatic means.

Shifts from graphics to text ‘Text has its uses.’ It is the general purpose tool of last resort. As Petre and Green observe, even the best-evolved notation will sometimes fail to express a given proposition; there must always be an escape from formalism, and the place of refuge usually lies in natural language. We have already noted that the cancellation of graphical implicatures can often be achieved via textual annotation. It is plausible to suggest that such cancellation is one of the major reasons why shifts from graphics to text should be necessary. Another relates to vocabulary size, to which we return shortly.

Viscosity ‘Massive changes to graphic representations are hard work . . . moving a piece of text is difficult . . . where connectivity has been established explicitly.’ This suggests that, at least in some cases, it is preferable to encode information pragmatically, rather than semantically, where there is a choice. This could be one reason why the diacritical re-inforcements considered by Marks and Reiter are *not* used in the real world schematics studied by Petre and Green.

Search trails ‘Searching is well supported in CAD-E in the domain of connectivity, but not in the domain of functionality.’ Petre and Green observe that symbol shape conveys a component’s function, but not the values it operates over. Such abstraction inevitably leads to “ambiguity”; moreover, a single physical component can have multiple graphical representations, giving differing views of its function. So violations of either brevity or informativeness are also a danger; as with natural language, any choice of a non-standard representation will appear meaningful, as a departure from the expected, basic-level choice. And the fake bi-stable flip-flop from Figure 8 shows that a constructed representation can accidentally resemble a basic-level entity.⁴ Petre and Green emphasise that, as with adjacency, correct choice of representations is an acquired skill, and that individual differences along this skill dimension tend to distinguish novice designers from experts. This situation is very similar to that confronting second language learners: pragmatic aspects of a language are often much harder to master than its semantics.

Vocabulary and space consumption ‘Most graphical notations rely to some extent on text to keep their vocabularies manageable.’ The problem of graphical ambiguity which we just noted is solved by using the additional modality of text: 50 to 65 component symbols are augmented with textual annotation to represent 100 to 500 physical part types. But why are graphical vocabularies kept so small? The answer lies in the domain of perceptual discriminability. It might not be impossible to design 500 discriminable symbols, but uncertainties over symbol interpretation would become more likely. It is better to say something general than to say something more precise, if the latter is destined to be misunderstood. The over-generality of the symbols is manageable because it is known that their full interpretation depends on the division of labour between graphics and language. In this way, *intended* ambiguity prevents unintended ambiguity.

⁴This possibility arises because basic-level entities can actually be graphically complex, resembling idiomatic phrases in natural language, rather than unanalysable words.

6.3 Comparison with cognitive dimensions

Petre and Green [1992:50] initially argue that the themes which emerged in their interviews are predictable from Fitter and Green's [1979] criteria, or from Green's [1989] cognitive dimensions framework. They ultimately suggest [pp65–7] that four aspects of the data are most important:

Information access Notations and their support tools should make the information conveyed as obvious as possible.

Escape from formalism Where tasks require information to be encoded that cannot be fitted into the main notational scheme, other mechanisms must be provided.

Secondary notation Users will exploit various techniques—such as informative layout—to go beyond the semantics of the official notation.

Individual differences Experts and novices differ particularly in their ability to use secondary notation effectively.

Fitter and Green's approach most squarely addresses the themes related to information access. There, it goes some way towards explaining why, for instance, graphics is good for overviews. However, because it is concerned with the usability and semantics of individual notations, the Fitter-Green approach does not touch on either the escape from formalism—possible only with multiple notations—or the pervasive use of secondary notation—possible only by going beyond the basic semantics of the notation. Nor were individual skill differences a matter of concern.

Green's cognitive dimensions, however, have a broader applicability. Because notations are seen in the context of task and support environment, the need to support the escape from formalism is explained. Secondary notation is useful because it reveals dependencies and encourages role-expressiveness. But because secondary notation is not automatically supported, some individuals will exploit it more effectively than others. Cognitive dimensions offer no direct explanation for limits on the graphical vocabulary; however, given the limits, cognitive dimensions do predict the necessity of a division of labour—some work carried out beyond the constraints of the graphical formalism. Similarly, various trade-offs can be predicted: premature commitment is acceptable if the overall system has low viscosity; or poor role-expressiveness in the main notation is permitted if support tools can reveal the roles using alternate notations.

How, then, does the Gricean account compare? Information access is at the core of the approach, and hence Fitter and Green is the closer model. Levinson [1983:121] noted that pragmatics should cover non-linguistic communication: 'If the maxims are derivable from considerations of rational co-operation, we should expect them to be universal in application, at least in co-operative kinds of interaction.' We have indeed seen that network designers appear to observe the Co-operative Principle. Their contributions are tailored to the purposes of their communications, and can be thought of as being carried out in accordance with Grice's maxims of conversation. Errors which frustrate information access can be seen as arising from violations of the maxims, which generate unwanted implicatures. The Gricean account

has greater reach than Fitter and Green, however. First, the escape from formalism is predicted to flow from the need to cancel unwanted implicatures of various kinds. Secondly, secondary notation is the very stuff of graphical pragmatics—meaningful structures which go beyond the plain semantics of the system. Finally, individual differences arise exactly because the mastery of pragmatic implicatures is one of the hardest things to achieve with a new representational system.

In this sense, the Gricean account goes beyond Fitter and Green, and covers territory that Petre and Green argue falls within the domain of cognitive dimensions. There is no question, however, of its replacing cognitive dimensions. The latter obviously has wider scope, taking into account tasks and support environments; unlike the implicature story, it can support claims concerning trade-offs between various system properties. For this reason, it is more likely to provide a useful guide to the design of human-computer interaction. However, we believe that the Gricean account complements cognitive dimensions, and is at least compatible with it. We have here concentrated on pragmatic aspects of the network notation, but pragmatic studies in general have always emphasised the need to look at notations in their context of use, and the crucial role of the notations' users. All discussion of pragmatic implicatures must ultimately be grounded in models of human inference, which show how we combine various goals and beliefs in context, to derive interpretations and actions. From this perspective, the Gricean account has many assumptions in common with cognitive dimensions. We look forward to constructive interaction between the two approaches.

7 Conclusions

We have argued here that there are certain parallels between pragmatic phenomena in NL and graphical representation. Elsewhere, formal techniques have been developed for modelling some of the NL phenomena. Thus, it is to be hoped that once the graphical data are better-understood, we may be able to treat them with formal techniques from NL pragmatics.

In the meantime, some might baulk at the idea of examining graphics from a perspective which was pioneered in linguistics, precisely because they attach value to real data, rather than to the toy examples so beloved of linguists. However, Levinson [1983:25] rightly observed that 'For many linguists, one of the major contributions of pragmatics has been to direct attention once again to *actual language usage*.' The discussion here, of Petre and Green's examples and observations, is intended to be only an initial step towards the pragmatic study of *actual graphical usage*.

References

- Austin, J. L. (1962) *How To Do Things With Words*. Oxford: Clarendon Press.
- Fitter, M. and Green, T. R. G. [1979] When do diagrams make good computer languages? *International Journal of Man-Machine Studies*, **11**, 235–261.
- Fodor, J. A. [1983] *Modularity of Mind*. Cambridge, Ma.: MIT Press.

- Gazdar, G. [1979] *Pragmatics: Implicature, Presupposition and Logical Form*. New York: Academic Press.
- Green, T. R. G. [1989] Cognitive dimensions of notations. In A. Sutcliffe and L. Macaulay (eds.) *People and Computers V*. Cambridge: Cambridge University Press.
- Grice, H. P. [1975] Logic and Conversation. In P. Cole and J. L. Morgan (eds.) *Syntax and Semantics 3: Speech Acts*, pp41–58. New York: Academic Press. Reprinted as Essay 2 in Grice [1989].
- Grice, H. P. [1978] Further Notes on Logic and Conversation. In P. Cole (ed.) *Syntax and Semantics 9: Pragmatics*, pp113–128. New York: Academic Press. Reprinted as Essay 3 in Grice [1989].
- Grice, H. P. [1989] *Studies in the Way of Words*. Cambridge, Ma.: Harvard University Press.
- Hirschberg, J. [1991] *A Theory of Scalar Implicature*. New York: Garland Publishing. Published version of 1985 thesis, University of Pennsylvania.
- Levinson, S. C. [1983] *Pragmatics*. Cambridge: Cambridge University Press.
- Marks, J. and Reiter, E. [1990] Avoiding unwanted conversational implicatures in text and graphics. In *Proceedings of the Eighth National Conference on Artificial Intelligence*, pp450–456. Boston, 29 July–3 August, 1990.
- Morris, C. W. [1938] *Foundations of the Theory of Signs*. In O. Neurath, R. Carnap and C. Morris (eds.) *International Encyclopedia of Unified Science*, pp77–138. Chicago: Chicago University Press.
- Petre, M. and Green, T. R. G. [1992] Requirements of graphical notations for professional users: electronics CAD systems as a case study. *Le Travail Humain*, **55**, 47–70.
- Rosch, E. [1978] Principles of Categorization. In E. Rosch and B. Lloyd (eds.) *Cognition and Categorization*, pp27–48. Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- Sadock, J. M. [1978] On testing for conversational implicature. In P. Cole (ed.) *Syntax and Semantics 9: Pragmatics*, pp281–298. New York: Academic Press.
- Searle, J. [1969] *Speech Acts: An Essay in the Philosophy of Language*. Cambridge: Cambridge University Press.
- Stalnaker, R. C. [1972] Pragmatics. In D. Davidson and G. Harman (eds.) *Semantics of Natural Language*, pp380–397. Dordrecht: D. Reidel.
- Thomason, R. H. [1977] Where pragmatics fits in. In A. Rogers, B. Wall and J. P. Murphy (eds.) *Proceedings of the Texas Conference on Performatives, Presuppositions and Implicatures*. Washington: Centre for Applied Linguistics.